International Journal of Computational Bioinformatics and In Silico Modeling
Vol. 3, No. 5 (2014): 502-513
Research Article
Open Access

aizeon publishers

ISSN: 2320-0634

In Silico analysis of Single Nucleotide Polymorphisms (SNPs) in human FANCA gene

Abubaker H. Mohamed[1], Ozaz Y. Mohammed[2], Sami AB. Salam[1], Hadeel A. Yousif[4]*, Mohamed M.Hassan[1], Hazim H. Kaheel[3] and Mohamed A.Hassan[1,6]

[1] Department of Bioinformatics, Africa city of Technology, Sudan
[2] Department of Hematology, University of Gezir, Sudan
[3] University, HNO –universities Klink-Tuebingen, Germany
[4] Faculty of Pharmacy, University of Khartoum, Sudan
[5] Division of Molecular Genetics, University of Tuebingen, Germany

*Corresponding author: Hadeel A. Yousif; e-mail: dolyyosef@yahoo.com

Received: 24 August 2014 Accepted: 18 September 2014 Online: 27 September 2014

ABSTRACT

Single-nucleotide polymorphisms (SNPs) play a major role in the understanding of the genetic basis of many complex human diseases. Also, the genetics of human phenotype variation could be understood by knowing the functions of these SNPs owing to the importance of FANCA gene in a post replication repair or a cell cycle checkpoint function. In this work, we have analyzed the genetic variation that can alter the expression and the function of the FANCA gene using computational methods. Genomic analysis of FANCA was initiated Polyphen and SIFT server used to retrieve 16 harmful mutations, among of these 16 nsSNPs damaged SNPs five non-synonymous SNPs showed very damaging by higher PSIC score of the Polyphen server with a SIFT tolerance index of 0.00-0.01 (R318M, I493T, A610T, P739L, R1117G). Protein structural analysis with these amino acid variants was performed by using I-Mutant and Modeling amino acid substitution with chimera software to check their stability and the effect of the native and mutant residues protein and structure for all 16 nsSNPs damaged. Screening for these SNPs variants in coding region may be useful for Fanconi anemia disease molecular diagnosis. Of the total 229 SNPs in 3′UTR region of FANCA gene, 24 SNPs were found in the 3′ UTR contain alleles can be disrupts a conserved miRNA site, therefore might change the protein expression levels.

Keywords: Single-nucleotide polymorphisms (SNPs), of FANCA gene, miRNA site, 3′UTR, Polyphen and SIFT

INTRODUCTION

Fanconi anemia (FA) is a rare inherited syndrome with diverse clinical symptoms including developmental defects, short stature, bone marrow failure, and a high risk of malignancies [1]. Fifteen genetic subtypes have been distinguished: FA-A, -B, -C, -D1, -D2, -E, -F, -G, -I, -J, -L, -M, -N, -O, and –P [2-5]. Since the first description by Guido Fanconi in 1927, FA has become to be recognized as an autosomal recessive disorder (X - linked in a rare subset) in which there is progressive bone marrow failure [6], as well there is an unusually high risk of developing treatment - resistant MDS and Acute Myeloid Leukemia AML, estimated at 52% in total by the age of 40 years. Furthermore, the risk of a variety of solid tumors, especially squamous cell carcinomas of the skin, is several times higher than in the general population [7]. Most, but not all, affected individuals also have one or more somatic abnormalities, including skin (caf é - au - lait spots), skeletal (absent thumbs, radial hypoplasia, scoliosis), genitourinary (underdeveloped gonads, horseshoe kidneys), gastrointestinal, cardiac and neurological anomalies.

The course of the disease and the pattern of somatic abnormalities show considerable variation, with approximately one - third of patients having no physical abnormalities. This makes diagnosis based on clinical criteria alone difficult and unreliable [6].

Cellular phenotype and function of the FA proteins

A wealth of recent genetic and biochemical data has shed new light on the elusive function of FA proteins. They constitute the FA pathway, a new DNA damage - response pathway responsible for repair of inter - and intra - strand DNA cross - links (ICL), either occurring spontaneously or induced by agents like mitomycin C or diexobutane .ICL lead to replication fork arrest and cell cycle arrest in late S phase. Completion of the cell cycle is not possible unless ICL are repaired [7]. These can result in increased numbers of chromosomal abnormalities including translocations and radial chromosomes. This hypersensitivity made treating FA patients a challenge in the past because traditional treatments of their symptoms resulted in more harm than good.[8],the most characteristic cellular feature of FA cells is the formation of DNA double - strand breaks on exposure to DNA inter – and intra - strand adducting agents (clastogens) such as mitomycin C and diexobutane. The in vitro response to clastogens has made it possible to test cells from different patients for their ability to cross - correct each other ' s defect by somatic cell fusion. As discussed earlier, this has led to classifying patients in 13 complementation groups.

The FA genes (genes that have been to be found mutated in FA patients) are called *FANC*, the most frequent being *FANCA*, *FANCC*, *FANCG*, and *FANCD2* [2].The corresponding *FANCA* gene is located on chromosome 16q24.3,this gene was reported to be involved in both inherited and somatic human diseases. The *FANCA* is the most frequently mutated gene in FA, representing with over 350 unique mutations reported Fanconi Mutation Database; (http://www.rockefeller.edu/fanconi/mutate),including large genomic deletions mediated by the unusually high density of ALU repetitions found in its genomic sequence [9]. For *FANCA*, associations of sequence alterations or altered expression have been suggested in some instances of ovarian cancer and leukemia [10, 11 and 12]

Mutations of FANCA account for about 60% of FA cases and are spread throughout the gene. None of the mutant alleles is common and few have been encountered more than once. Thus, identifying mutations in newly diagnosed cases of FA is laborious, as one needs to scan the entire coding sequence (of several genes, unless the complementation group is known) [7], most of the *FANCA*-activating non-synonymous mutations are distributed in c.3788 3790del (p.Phe1263del), c.1115 1118delTTGG (p.Val372fs) of European and Brazilian population, Exon 12–17del and Exon 12–31del of South-Africans, and c.295C>T (p.Gln99X) of Spanish Gypsy population[9, 13-15].

SNPs stand for Single Nucleotide Polymorphisms are DNA sequence variations that occur when a single nucleotide (A, T, C, or G) in the genome is altered. They are found throughout the genome in exons, introns intergenic regions, promoters, enhancers[16], SNP in a promoter can influence gene expression [16], and thus more likely to contain an allele being more functionally or physiological relevant than other types of polymorphism.

The identification of SNPs responsible for specific phenotypes seems to be a problem, since requiring multiple testing of hundreds or thousands of SNPs in candidate genes [17]. Recently researchers found new functional polymorphisms called MiRSNPs/ polymorphism located at micro RNA binding sites of functional gene that can influence gene expression by interfering with microRNA function of those SNPs within microRNAs (miRNAs) [18].

Inour study the decision of choosing the right set of SNPs from the NCBI Human reference genome (www.ncbi.nlm.nih.gov), to be screened was a critical one. One possible way to overcome this problem was to prioritize SNPs according to their structural and functional significance using different bioinformatics prediction tools.

An important role of the FA pathway in the biology of sporadic cancers is increasingly being recognized. Epigenetic silencing of FAC genes has been identified in different solid tumors and acute leukemias and in some cases correlated with their sensitivity to chemotherapeutic agents causing DNA cross - links such as cisplatin, cyclophosphamide and melphalan. In a small proportion of patients with AML, somatic mutations in FANCA were identified but their functional significance is unknown. Therefore, studying the functional status of the FA pathway in different tumors may offer valuable information on which therapeutic choices can be based [7].

MATERIALS AND METHODS

The data on human FANCA genewas collected by NCBI Genome Workbench software version 2.7.6 and Entrez Gene on Search National Center for Biological Information Public Databases. The SNPs information (Protein accession number and SNP ID) of the FANCA gene was retrieved from the NCBI dbSNP (http://www.ncbi.nlm.nih.gov/snp/), and SWISSProt databases (http://expasy.org/).

SAAP and dbSNPs

Single Amino Acid Polymorphisms (SAAP) (http://www.bioinf.org.uk/saap/) [19] and Single Nucleotide Polymorphism Database (dbSNPs)were used to recognize the single point mutation within the encoded protein of *FANCA*. In order to check the effect of the mutation on the protein (damaging or not) we used the SIFT-software (online) upon which by applying the Polyphen software the previous SIFT-results were re-confirmed.

PolyPhen

PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) is an online bioinformatics program to automatically predict the consequence of an amino acid change on the structure and function of a protein. This prediction

is based on a number of features comprising the sequence, phylogenetic and structural information characterizing the substitution. Basically, this program searches for 3D protein structures, multiple alignments of homologous sequences and amino acid contact information in several protein structure databases, then calculates position-specific independent count scores (PSIC) for each of two variants, and then computes the PSIC scores difference between two variants. The higher a PSIC score difference, the higher the functional impact a particular amino acid substitution is likely to have PolyPhen scores were assigned probably damaging (2.00 or more), possibly damaging (1.40–1.90), potentially damaging (1.0–1.50), benign (0.00–0.90). BasicallyPolyPhen accepts input in form of SNPs or protein sequences [20].

SIFT

In order to detect deleterious nsSNPs (non-synonymous) program version2 (http://blocks.fhcrc.org/sift/SIFT.html) was used, SIFT is a novel bioinformatics tool to predict whether an amino acid substitution affects protein function, so that users can prioritize substitutions for further study. The main underlying principle of this program is that it generates alignments with a large number of homologous sequences, and assigns scores to each residue ranging from zero to one. Scores close to zero indicate evolutionary conservation of the genes and intolerance to substitution, while scores close to one indicate tolerance to substitution only [21].

I-Mutant

We used I-Mutant version 3.0 to predict the protein stability changes upon single-site mutations. I-Mutant3.0 basically can evaluate the stability change of a single site mutation starting from the protein structure or from the protein sequences. This program was trained on some data set derived from ProTherm[22] which is considered to be the most comprehensive database of experimental data on protein mutations.

Modelling

Modeling the SNPs on the 3D structure of the proteins is a very helpful action in order to predict the impact of SNPs on structural level. Therefore we used CPHmodels 3.2 server to predict the 3D model for those proteins with an unknown 3D structure [23], the protein sequence used were obtained from the ExPASy Database (www.expasy.org/).

Chimera

Chimera (http://www.cgl.ucsf.edu/chimera) is a high-quality extensible program for interactive conception and analysis of molecular assemblies and related data [24].Chimera (version 1.6.2), this software was produced by University of California, San Francisco (UCSF),we used in this step to generate the mutated models of each *FANCA*protein 3D model entries. This software is used to browse respectively locate the 3D structure of the specific protein and then alter the native amino acid with a mutated one and to look for structural effect that may be produced. The outcome is then a graphic model depicting the mutation.

PolymiRTS

(http://compbio.uthsc.edu/miRSNP/) is the database server designed specifically for the analysis of the 3'UTR region, at this stage we used this server to determine SNPs that may alter miRNA target sites. All SNPs located within the 3′-UTRs of database were selected separately and submitted to the program. Then we checked if the SNP variants could alter putative miRNA target sites [25], focusing on mutations that alter sequence complementarity to miRNA seed regions possibly leading to fanconi anemia.

RESULTS AND DISCUSSION

Retrieval of SNPs

The *FANCA* gene investigated in this work were retrieved from the dbSNP National Center of Biotechnology Information (NCBI public database) Genome workbench software. *FANCA* gene contained a total of 3332 SNPs, of which 204 were non synonymous SNPs, while in *3'UTR*region were 229SNPs, total SNPs contain SNPs of *3'UTR/5'UTR* near to gene, non-coding *3'UTR/5'UTR*, intron, coding synonymous and coding non-synonymous regions. We selected non-synonymous coding SNPs and *3'UTR* SNPs for our investigation.

Deleterious or damaging nsSNPs predicted by PolyPhen and SIFT

The RS number (reference SNPs) of 204 nsSNPs *FANCA* gene were submitted cluster as batch to the SIFT and PolyPhen servers. Among the 204 nsSNPs, 16 SNP were predicted found to be damaging and deleterious by both the SIFT and PolyPhen servers, the polyphen program have position-specific independent count (PSIC) score above of 1.0 is considered to be damaging. The results are shown in (Table 1), H186Q, L193V, F319L, F836Y and S850Y, showed possibly damaging prediction by PolyPhen servers, whereas S290Y, R318M, I493T, A610T, C625S, S674L, P739L, R1117G, R1144W and T1161K showed probably damaging for score higher than 0.95. The SIFT program result showed all these mutations were damaging with a tolerance index score of ≤0.05 as deleterious, one nsSNPs showed a highly deleterious tolerance index score of 0.00 is P739L (Table 1), this variant was showed reported in swiss-variant (www.expasy.org/) database but not specific disease.R1117G mutation was earlier reported for their role in Fanconi anemia[13, 26]. Given the fact that nsSNPs in critical cellular genes such as FANCA modify the DNA repair protein that may operate in a postreplication repair or a cell cycle checkpoint function it is believed to play an important role in disease predisposition also. Therefore, an effort was made to identify SNPs that can modify the structure, function and expression of the FANCA gene.

Table 1. Prediction result of SIFT and Polyphen programs

Gene name	SNP ID	Chromosome location	Nucleotide change	Amino acid Change	PolyPhen-2 result	PSIC SD	SIFT result	Tolerance Index
FANCA	rs149159377	chr16,89874740	C/G	H186Q	possibly damaging	1.098	DAMAGING	0.01
	rs141861208	chr16,89874721	C/G	L193V	possibly damaging	1.502	DAMAGING	0.02
	rs185984960	chr16,89865598	C/A	S290Y	probably damaging	2.478	DAMAGING	0.05
	rs72552377	chr16,89862367	G/T	R318M	probably damaging	2.803	DAMAGING	0.01
	rs17226068	chr16,89862363	T/C	F319L	possibly damaging	1.691	DAMAGING	0.01
	rs113042003	chr16,89849503	T/C	I493T	probably damaging	1.852	DAMAGING	0.01
	rs192541465	chr16,89842222	G/A	A610T	probably damaging	2.022	DAMAGING	0.01
	rs146491000	chr16,89842179	C/G	A624G	probably damaging	1.743	DAMAGING	0.03
	rs139235751	chr16,89842176	T/A	C625S	probably damaging	2.034	DAMAGING	0.02
	rs17232973	chr16,89838216	C/T	S674L	probably damaging	2.146	DAMAGING	0.04
	rs45441106	chr16,89836978	C/T	P739L	probably damaging	2.533	DAMAGING	0
	rs142869950	chr16,89833643	T/A	F836Y	possibly damaging	1.893	DAMAGING	0.02
	rs77805547	chr16,89833601	C/A	S850Y	possibly damaging	2.5	DAMAGING	0.02
	rs149277003	chr16,89813298	A/G	R1117G	probably damaging	1.866	DAMAGING	0.01
	rs143671872	chr16,89813075	C/T	R1144W	probably damaging	3.464	DAMAGING	0.03
	rs142833057	chr16,89813023	C/A	T1161K	probably damaging	1.902	DAMAGING	0.02

***PolyPhen-2 result**: POROBABLY DAMAGING (more confident prediction) / POSSIBLY DAMAGING (less confident prediction), PSIC SD: Position-Specific Independent Counts software if the score is =>0.5 ,Tolerance Index: Ranges from 0 to 1. The amino acid substitution is predicted damaging if the score is <= 0.05, and tolerated if the score is > 0.05.*

Table 2. Prediction result of I-Mutant software

RI	DDG Value Prediction Kcal/mol	SVM2 Prediction Effect	Temp (°C)	PH	MT	WT	Amino acid position	SNP ID	Gene name
1	0.11	Decrease	25	7	Q	H	186	rs149159377	**FANCA**
4	-1.19	Decrease	25	7	V	L	193	rs141861208	
1	0.07	Increase	25	7	Y	S	290	rs185984960	
4	-0.52	Decrease	25	7	M	R	318	rs72552377	
5	-1.31	Decrease	25	7	L	F	319	rs17226068	
8	-2.05	Decrease	25	7	T	I	493	rs113042003	
6	-0.71	Decrease	25	7	T	A	610	rs192541465	
8	-1.2	Decrease	25	7	G	A	624	rs146491000	
7	-0.76	Decrease	25	7	S	C	625	rs139235751	
0	-0.11	Increase	25	7	L	P	739	rs45441106	
2	0.11	Increase	25	7	L	S	674	rs17232973	
2	-0.91	Decrease	25	7	Y	F	836	rs142869950	
0	-0.36	Increase	25	7	Y	S	850	rs77805547	
8	-1.57	Decrease	25	7	G	R	1117	rs149277003	
6	-0.52	Decrease	25	7	W	R	1144	rs143671872	
6	-0.85	Decrease	25	7	K	T	1161	rs142833057	

Prediction of change in stability due to mutation used I-Mutant 3.0 server

I mutant 3.0 results demonstrated that protein stability with related free energy had changed due to mutation. The twelve mutation (H→Q, L→V, R→M, F→L, I→T, A→T, A→G, C→S, F→Y, R→G, R→W)in *FANCA* gene that decrease effective stability of the protein (Table 2), and four mutation(S→Y, P→L, S→L, S→Y) that increase effect stability.

Modeling of amino acid substitution effects due to nsSNPs on protein structure

FANCA Protein sequences of the nsSNP were presented to CPH 3.2 server to get the model sequence as PDB file. After that Chimera program had been to visualize the PDB file, determine the position of the mutant and replaced it with the new amino acid, which showed based structural change among nsSNPs protein and clash points of new residues with other atoms using minimize energy. Each amino acid in wild type and mutant structures carries specific properties like solvent accessibility, charge density, hydrophobicity, rigidity, molecular surfaces and electrostatic potential values. Native and mutant residues sometimes differ due to carrying such specific properties and can disrupt the structural and functional features of the original protein. All native and mutant structure of *FANCA* protein showed in the (Fig. 1, see supplementary data).

The c.558C>G mutation (rs149159377), which leads to His→Gln conversion at 186 position of *FANCA* protein modeling (Fig 1.A) is significant because the basic positive amino acid is transformed into negatively charge acidic amino acid that's possibly disturbs the electrostatic interactions with other molecules, this mutation can form hydrogen bond and contact with other Gln220 and Cys216 nearby are 1.867 Å and 1.523Å respectively.

The c.869C>A mutation (rs185984960) that leads to Ser→Tyr at 290 position observed variant due to the transformation of sulfur into aromatic neutral amino acid that can disrupts the proper folding of *FANCA* protein complex (Fig 1.C). The mutated Tyr290 hydrogen bonding interaction with nearby Trp346, His292 and Leu362.

The c.1478T>C mutation (rs113042003) that leads to Ile→Thr conversion at 493 position is different from other mutations because the hydrophobic neutral amino acid is transformed into hydrophilic amino acid with hydroxyl group (Fig 1.F). The mutated Thr493induces different conformational changes and distort the protein stability due to clashes with H-bonding 1.534 Å nearby of Glu542.

The c.2216C>T variant (rs45441106) that leads to Pro→Leu conversion at 739 a position is a rarely observed variant due to the transformation of cyclic amino acid into chain residue (Fig 1.K). Prolines are known for their rigidity and therefore create special backbone conformation, which might be disturbed and clashes at this position due to Leucine, this current mutation can affect the structure and function of native structure due to carrying the less charge density, aliphatic domain and less rigidity.

Thec.3430C>T variant (rs143671872) leads to Arg→Trp conversion at position 1144, this variant mutated residue is found to be bigger because contain tow cycle aromatic group, neutral and more hydrophobic than the wild type, therefore consider as probably damaging disturbed structure and function of *FANCA* protein, and showed clashes with more than one residue surrounding in the structure (Fig. 1.O)

Functional SNPs in 3 untranslated regions (UTR) predicted by PolymiRTS data base 3.0

SNPs in *3′UTR* of *FANCA* gene were presented to PolymiRTS server, output shown, Among 229 SNPs in *3′UTR* region of *FANCA* gene 24 functional SNPs was predicted, through the 24 SNPs, 19 allele disrupts a conserved of76 miRNA site (ancestral allele with support >= 2), 21 derived allele creates a new site of 73 miRNA (Table 3, see supplementary data), rs200396402 SNPs contain (C) allele have 12miRSite as target binding site can be disrupts a conserved miRNA site.

Regarding the fact that FA patients at high risk to developed other malignancies, so early and specific diagnosis can lead to better prognosis. The chromosome breakage test yields occasional false positives, as other genetic disorders, such as Nijmegen breakage syndrome and Roberts syndrome, also display aberrant chromosomes upon exposure to these DNA interstrand cross-linking DNA ICLs [27, 28]. Patients with fanconi anemia have same clinical and cellular characteristic with some exceptions, this mean that the 15 genes work together in chronological order to maintain normal function of the fanconi anemia pathway, so any abnormality defected in this pathway will result in DNA repair defect and cell cycle which can lead also to the genetic in stability in other somatic cell. As evidence many defects of the FA genes have been found in a wide variety of human cancers in the general population (patients without FA) [29, 30]. Also defects in the FA pathway may be responsible for the hypersensitivity of cancer cells to certain types of chemotherapeutic drugs and radiation. Disruption of FA genes may become a useful predictor of sensitivity to chemotherapy with widely used anticancer DNA cross-linking agents (cisplatin, MMC and melphalan [29, 31, 32].

CONCLUSION

The *FANCA* gene was investigated in this work by evaluating the influence of functional SNPs through computation methods, out of a total of 3332 SNPs in the *FANCA*, 204 were found to be non-synonymous and 229 were found to be in the 3′untranslated regions, 16 nsSNPs were found to be deleterious and damaging by both Polyphen and SIFT server, 24 SNPs in the 3′ UTR were found to be of functional significance. In order to make effective use of genetic diagnosis, the harm SNPs in all fanconi genes should be well known and available to the diagnostic services and molecular biology laboratories to ensure accurate diagnosis for this complicated disease which can also lead to successful intervention which dependent on finding the cause or causes of a problem.

From these results we can conclude the *FANCA* gene is very important in the diagnosis of mutation caused by Fanconi anemia disease that's contain big number of damaging SNPsby coding non-synonymous and 3′UTR regions.

Acknowledgments

The authors wish to acknowledge the enthusiastic cooperation of Africa City of Technology, Sudan.

Competing interests

The authors declare that they have no competing interests.

REFERENCES

1. Anneke B. Oostra, Aggie W. M. Nieuwint, Hans Joenje, et al. (2012). Diagnosis of Fanconi Anemia: Chromosomal Breakage Analysis. Anemia. Vol. 2012
2. deWinter J. P. and Joenje H. (2009). The genetic and molecular basis of Fanconi anemia, Mutation Research 668 (1-2) 11–19
3. Vaz F. H. Hanenberg, B. Schuster et al. (2010). Mutation of the RAD51C gene in a Fanconi anemia-like disorder. Nature Genetics, 42 (5): 406–409.
4. Stoepker C., K.Hain, B. Schuster et al. (2011). SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. Nature Genetics 43 (2): 138–141.
5. Kim Y., F. P. Lach, R. Desetty et al. (2011). Mutations of the SLX4 gene in Fanconi anemia Nature Genetics 43 (2): 142–146.
6. A Victor Hoffbrand, Daniel Catovsky, Edward GD Tuddenham et al. (2011). Sixth edition, chapter 12, page 186-189 Blackwell Publishing Ltd
7. Drew Provan, John G. Gribben, (2010). Molecular Hematology, 3rd ed, chapter 12, page 153-154 ISBN 978-1-4051-8231-7, Blackwell Publishing Ltd.
8. AllisonM. Green and GaryM. Kupfer (2009). Fanconi Anemia Hematol Oncol Clin N Am 23:193–214
9. Levran O., Diotti R., Pujara K. et al. (2005). Spectrum of sequence variations in the FANCA gene: an International

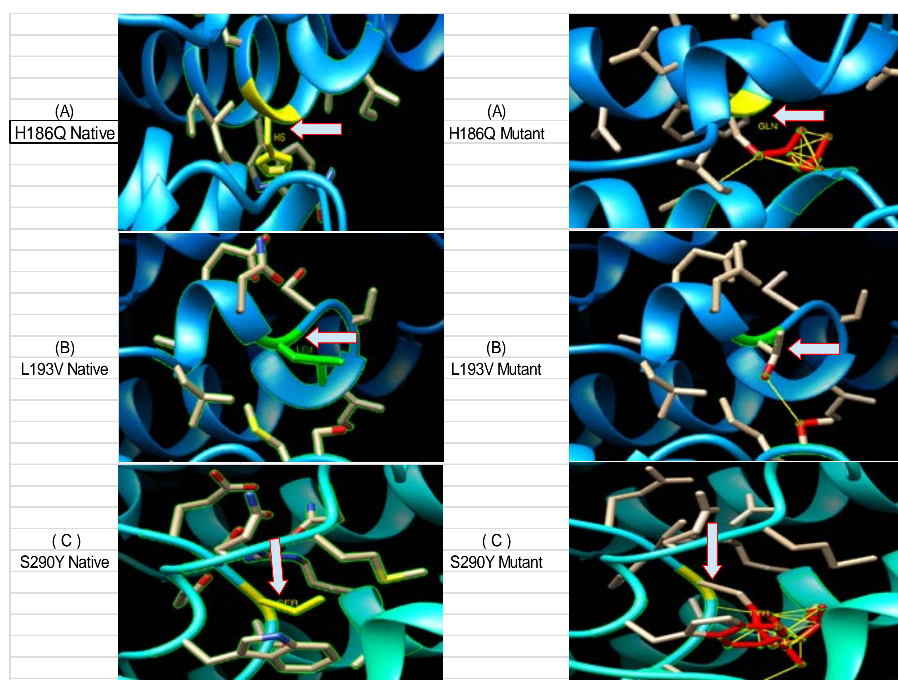
Fanconi anemia Registry (IFAR) study. Hum. Mutat. 25, 142e149.
10. Thompson E., Dragovic R.L., Stephenson S.A. et al. (2005). A novel duplication polymorphism in the FANCA promoter and its association with breast and ovarian cancer. BMC Cancer 5, 43.
11. Lensch M.W., Tischkowitz M., Christianson T.A., et al. (2003). Acquired FANCA dysfunction and cytogenetic instability in adult acute myelogenous leukemia. Blood 102, 7e16.
12. Tischkowitz M., Easton D.F., Ball J. et al. (2008). Cancer incidence in relatives of British Fanconianaemia patients. BMC Cancer 8, 257.
13. Levran O., T. Erlich, N. Magdalena et al. (1997). Sequence variation in the Fanconi anemia gene FAA. Proceedings of the National Academy of Sciences of the United States of America 94 (24): 13051–13056.
14. Magdalena N., D. V. Pilonetto, M. A. Bitencourt et al., (2005). Frequency of Fanconi anemia in Brazil and efficacy of screening for the FANCA 3788-3790del mutation. Brazilian Journal of Medical and Biological Research 38 (5): 669–673.
15. Tipping A. J., T. Pearson, N. V. Morgan et al., (2001). Molecular and genealogical evidence for a founder effect in Fanconi anemia families of the Afrikaner population of South Africa. Proceedings of the National Academy of Sciences of the United States of America, 98 (10): 5734–5739.
16. Drazen JM, Yandava CN, Dube L et al. (1999). Pharmacogenetic association between ALOX5 promoter genotype and the response to anti-asthma treatment. Nat Genet 22: 168–170.
17. Krawczak, E. V. Ball, I. Fenton (2000). Human gene mutation database-a biomedical information and research resource. Human Mutation, 15(1): 45–51.
18. Sethupathy P. and Collins F.S. (2008). MicroRNA target site polymorphisms and human disease. Trends Genet. 24: 489–497.
19. Cavallo A. and Martin A.C., (2005). Mapping SNPs to protein sequence and structure data. Bioinformatics 8: 1443–1450.
20. Ramensky, V., Bork, P., Sunyaev, S., (2002). Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 30: 3894–3900.
21. Ng, P.C., Henikoff, S. (2001). Predicting deleterious amino acid substitutions. Genome Res. 11(5): 863-74.
22. Capriotti E., Calabrese R. and Casadio R. (2006). Predicting the insurgence of human genetic disease associated to single point protein mutation with Support Vector Machines and evolutionary information. Bioinformatics 22: 2729-2734.
23. Nielsen M., Lundegaard C., Lund O., et al. (2010). CPHmodel-3.0-Remote homology modeling using structure guided seguence profiles. Nucleic Acids Research Vol.38, doi: 10.1093/nar/gkq535.
24. Pettersen EF, Goddard TD, Huang CC, et al. (2004). Comput UCSF Chimera--a visualization system for exploratory research and analysis. Chem. 25(13): 1605-12.
25. Bhattacharya A. Ziebath JD and Cui Y. (2014). PolmiRTS Database 3.0: Linking polymorphisms in microRNAs and their target sites with human diseases and biological pathway. Nucleic Acids Res. 42(D1): D86-D91.
26. Kupfer G., Naef D., Garica-Higera I. et al. (1999). A patient-derived mutant from of the fanconi anemia protein, FANCA, is defective in nuclear accumulation. Exp. Hematol. 27:587-593.
27. Younghoon Kee and Alan D. D'Andrea (2012). Molecular pathogenesis and clinical management of Fanconi anemia. J Clin Invest. 122(11): 3799–3806.
28. Van der Lelij P, Oostra AB, Rooimans MA, et al. (2010). Diagnostic overlap between Fanconi anemia and the Cohesinopathies: Roberts Syndrome and Warsaw Breakage Syndrome. Anemia. 2010:565268.
29. Taniguchi T, Tischkowitz M, Ameziane N, et al. (2003). Disruption of the Fanconi anemia-BRCA pathway in cisplatin-sensitive ovarian tumors. Nat Med. 9:568-574
30. Turner N, Tutt A and Ashworth A. (2004). Hallmarks of 'BRCAness' in sporadic cancers. Nat Rev Cancer. 4:814-819.
31. Van der Heijden MS, Brody JR, Gallmeier E, et al. (2004). Functional defects in the Fanconianemia pathway in pancreatic cancer cells. Am J Pathol. 165:651-657
32. Chen Q, Van der Sluis PC, Boulware D. et al. (2005). The FA/BRCA pathway is involved in melphalan-induced DNA interstrand cross-link repair and accounts for melphalan resistance in multiple myeloma cells. Blood. 106:698-705

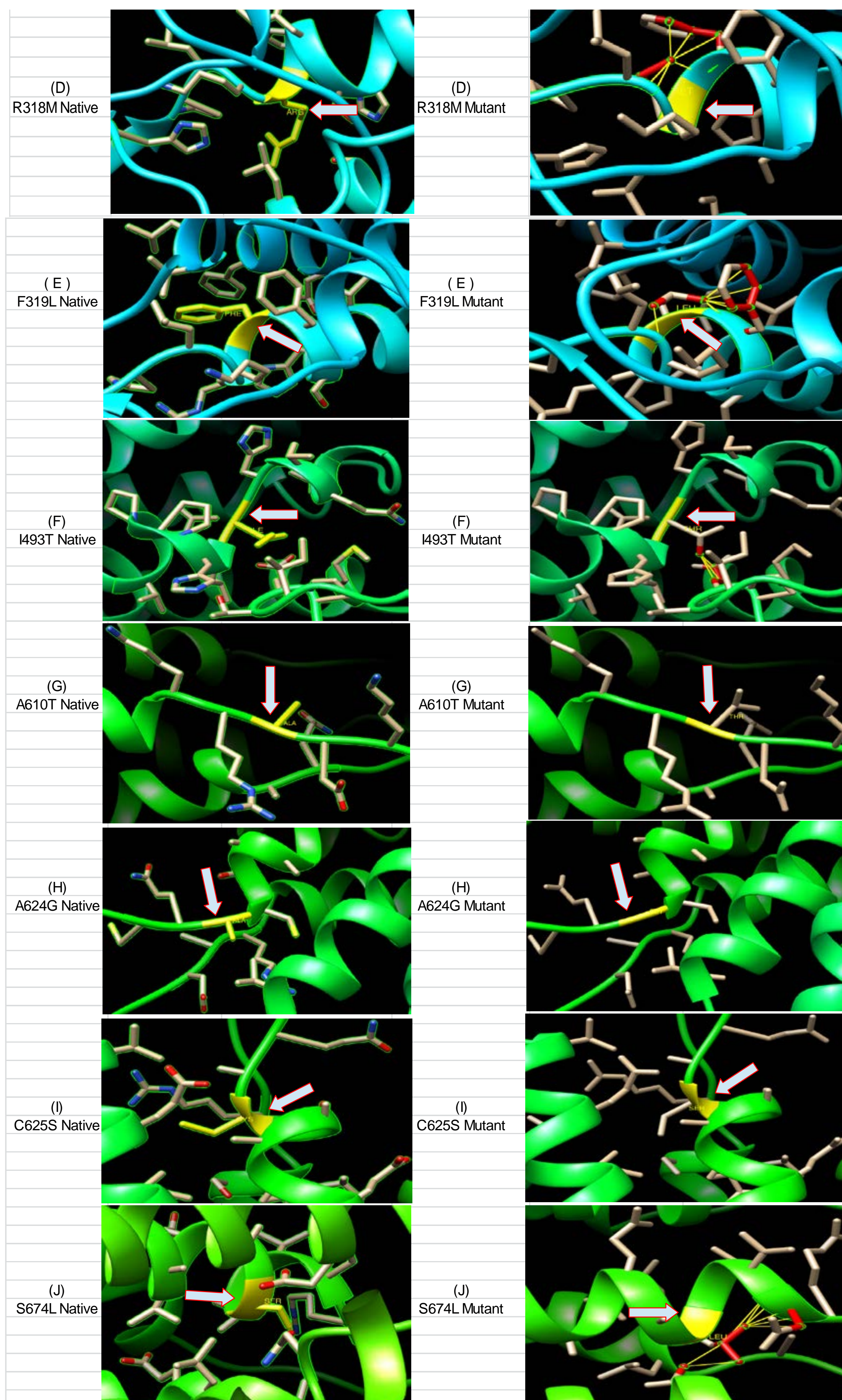
© 2014; AIZEON Publishers; All Rights Reserved

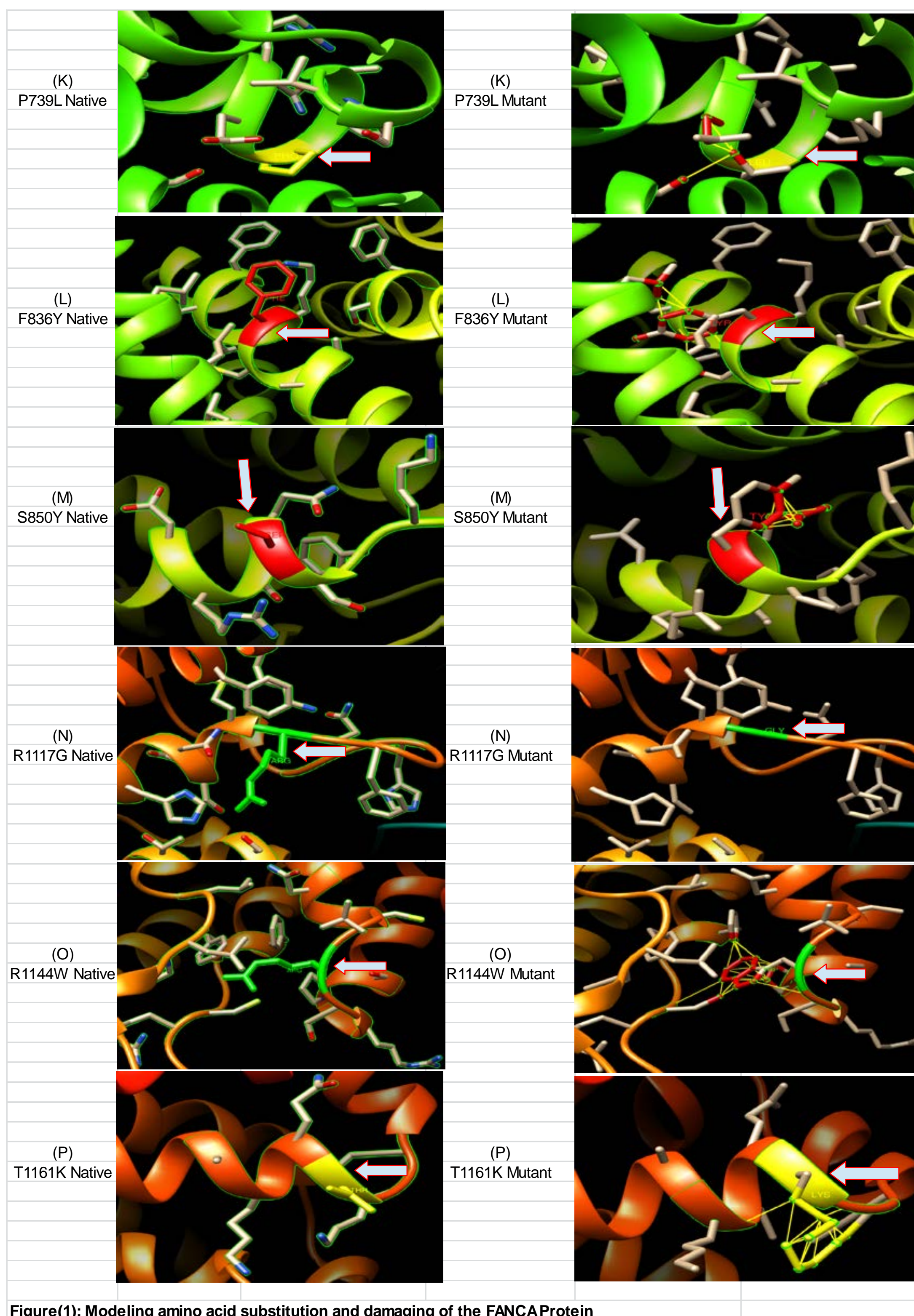
This is an Open Access article distributed under the terms of the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*****

Supplementary Data




(D)
R318M Native
(D)
R318M Mutant
(E)
F319L Native
(E)
F319L Mutant
(F)
I493T Native
(F)
I493T Mutant
(G)
A610T Native
(G)
A610T Mutant
(H)
A624G Native
(H)
A624G Mutant
(I)
C625S Native
(I)
C625S Mutant
(J)
S674L Native
(J)
S674L Mutant



Figure(1): Modeling amino acid substitution and damaging of the FANCA Protein

Table 3. SNPs and INDELs in miRNA target sites at 3′-UTRs region of FANCA gene

Function class	miRSite	Conservation	miR ID	dbSNP ID
D	gCATTTCAggatg	2	hsa-miR-203a	rs17227452
C	gCATTTTAggatg	2	hsa-miR-3662	
				rs181463391
C	tttataTACTGTC	2	hsa-miR-4645-3p	
D	caaccTCCCGGGg	2	hsa-miR-1247-3p	rs201316239
D	CAACCTCccgggg	2	hsa-miR-377-5p	
D	caacctCCCGGGG	2	hsa-miR-5587-3p	
D	CAACCTCccgggg	2	hsa-miR-6086	
D	caACCTCCCGggg	2	hsa-miR-6799-5p	
C	caaCCTTCCGGgg	2	hsa-miR-6125	
D	gtgagTCCAGGCc	2	hsa-miR-1254	rs200396402
D	gtgagTCCAGGCc	2	hsa-miR-3116	
D	gtgaGTCCAGGcc	2	hsa-miR-3690	
D	gtgAGTCCAGgcc	2	hsa-miR-378a-3p	
D	gtgAGTCCAGgcc	2	hsa-miR-378b	
D	gtgAGTCCAGgcc	2	hsa-miR-378c	
D	gtgAGTCCAGgcc	2	hsa-miR-378d	
D	gtgAGTCCAGgcc	2	hsa-miR-378e	
D	gtgAGTCCAGgcc	2	hsa-miR-378f	
D	gtgAGTCCAGgcc	2	hsa-miR-378h	
D	gtgAGTCCAGgcc	2	hsa-miR-378i	
D	gtgAGTCCAGgcc	2	hsa-miR-422a	
C	gtgAGTGCAGgcc	2	hsa-miR-4793-3p	
C	gtGAGTGCAggcc	2	hsa-miR-7703	
D	agtgagTCCAGGC	2	hsa-miR-1254	rs186050933
D	agtgagTCCAGGC	2	hsa-miR-3116	
D	AGTGAGTccaggc	2	hsa-miR-3680-5p	
D	agtgaGTCCAGGc	2	hsa-miR-3690	
D	agtgAGTCCAGgc	2	hsa-miR-378a-3p	
D	agtgAGTCCAGgc	2	hsa-miR-378b	
D	agtgAGTCCAGgc	2	hsa-miR-378c	
D	agtgAGTCCAGgc	2	hsa-miR-378d	
D	agtgAGTCCAGgc	2	hsa-miR-378e	
D	agtgAGTCCAGgc	2	hsa-miR-378f	
D	agtgAGTCCAGgc	2	hsa-miR-378h	
D	agtgAGTCCAGgc	2	hsa-miR-378i	
D	agtgAGTCCAGgc	2	hsa-miR-422a	
C	agtgAGCCCAGgc	2	hsa-miR-378g	
C	aGTGAGCCcaggc	2	hsa-miR-4284	
C	agTGAGCCCAggc	2	hsa-miR-4437	
C	agtGAGCCCAggc	2	hsa-miR-4674	
C	agtgagCCCAGGC	2	hsa-miR-661	
C	agtgaGCCCAGGc	2	hsa-miR-939-3p	
D	aGAGGGGAgagtc	7	hsa-miR-6795-3p	rs199778951
D	aGAGGGGAgagtc	7	hsa-miR-6826-3p	
D	aGAGGGGAgagtc	7	hsa-miR-6887-3p	

C	aGAGGGTAgagtc	7	hsa-miR-7848-3p	
C	acattcATCACAG	8	hsa-miR-4679	rs138414698
D	gcttcCGCTGCTt	13	hsa-miR-503-5p	
C	gcttCCCCTGCtt	9	hsa-miR-1909-3p	
C	gCTTCCCCtgctt	10	hsa-miR-4271	
C	gCTTCCCCtgctt	10	hsa-miR-4725-3p	
C	gcttcCCCTGCTt	13	hsa-miR-6510-5p	rs149209309
C	gcttCCCCTGCtt	9	hsa-miR-6722-3p	
C	gCTTCCCCtgctt	10	hsa-miR-6780b-5p	
C	gcTTCCCCTgctt	12	hsa-miR-6783-5p	
C	CCCTGGAaggggg	2	hsa-miR-4685-5p	
C	CCCTGGAaggggg	2	hsa-miR-6837-5p	rs17233819
C	CCCTGGAaggggg	2	hsa-miR-7113-5p	
D	aGCCCTGGgaggg	2	hsa-miR-4685-5p	
D	aGCCCTGGgaggg	2	hsa-miR-6837-5p	rs17233812
D	agggccACAGCCC	22	hsa-miR-4675	
D	agggccACAGCCC	22	hsa-miR-4741	
D	agGGCCACAgccc	2	hsa-miR-7108-5p	
D	agggcCACAGCCc	5	hsa-miR-7974	
C	agggccCCAGCCC	22	hsa-miR-4492	
C	agggccCCAGCCC	22	hsa-miR-4498	
C	agggcCCCAGCCc	5	hsa-miR-4505	rs4399534
C	agggCCCCAGCcc	5	hsa-miR-4731-5p	
C	agggccCCAGCCC	22	hsa-miR-5001-5p	
C	agggcCCCAGCCc	5	hsa-miR-5787	
C	agGGCCCCAgccc	2	hsa-miR-6762-5p	
C	agGGCCCCAgccc	2	hsa-miR-6845-5p	
C	agggccCCAGCCC	22	hsa-miR-762	
D	cgagGGCCACAgc	2	hsa-miR-7108-5p	rs144483660
C	cgagGGTCACAgc	2	hsa-miR-6808-3p	
D	gtcaTGTGGTAct	18	hsa-miR-140-3p	
D	gtcatgTGGTACT	18	hsa-miR-624-5p	rs201383247
C	GTCATGCggtact	20	hsa-miR-6076	
C	GTCATGCggtact	20	hsa-miR-6797-3p	
D	cTCAGCCTtgtgt	8	hsa-miR-3929	
D	cTCAGCCTtgtgt	8	hsa-miR-4419b	
D	ctcagCCTTGTGt	12	hsa-miR-4457	
D	cTCAGCCTtgtgt	8	hsa-miR-4478	
D	ctcagcCTTGTGT	8	hsa-miR-4731-3p	
D	ctcagcCTTGTGT	8	hsa-miR-4801	rs142475962
C	ctCAGCTTTgtgt	8	hsa-miR-320a	
C	ctCAGCTTTgtgt	8	hsa-miR-320b	
C	ctCAGCTTTgtgt	8	hsa-miR-320c	
C	ctCAGCTTTgtgt	8	hsa-miR-320d	
C	ctCAGCTTTgtgt	8	hsa-miR-4429	
D	cgCCGGCCAcact	10	hsa-miR-4743-5p	rs144644592

D	cgccgGCCACACt	10	hsa-miR-644a	
D	cgccGGCCACAct	10	hsa-miR-663b	
D	cgccGGCCACAct	10	hsa-miR-7108-5p	
C	cgccGGTCACAct	10	hsa-miR-6808-3p	
C	cgccGGTCACAct	10	hsa-miR-7973	
C	gtgcaCGATGGAc	15	hsa-miR-181a-3p	
C	gTGCACGAtggac	9	hsa-miR-4536-3p	
C	gTGCACGAtggac	9	hsa-miR-517a-3p	rs200903461
C	gTGCACGAtggac	9	hsa-miR-517b-3p	
C	gTGCACGAtggac	9	hsa-miR-517c-3p	
D	cggTGGCCCAggt	2	hsa-miR-4640-5p	
D	cggTGGCCCAggt	2	hsa-miR-4726-5p	
D	cggTGGCCCAggt	2	hsa-miR-6508-3p	
D	cggtggCCCAGGT	4	hsa-miR-7851-3p	rs201151160
D	cggtgGCCCAGGt	5	hsa-miR-939-3p	
C	cggTGGGCCAggt	2	hsa-miR-3136-3p	
C	cggTGGGCCAggt	2	hsa-miR-7155-3p	
D	cacagaGGGGCTC	2	hsa-miR-1225-3p	
D	caCAGAGGGgctc	4	hsa-miR-2276-5p	
D	cacaGAGGGGCtc	4	hsa-miR-4749-3p	
D	cacagAGGGGCTc	4	hsa-miR-6769a-3p	
C	cacagaAGGGCTC	2	hsa-miR-1233-3p	
C	caCAGAAGGgctc	4	hsa-miR-1237-3p	rs145686036
C	cacagAAGGGCTc	4	hsa-miR-129-1-3p	
C	cacagAAGGGCTc	4	hsa-miR-129-2-3p	
C	cACAGAAGgctc	2	hsa-miR-3182	
C	cacAGAAGGGctc	4	hsa-miR-5088-3p	
C	cacaGAAGGGCtc	4	hsa-miR-6886-3p	
D	gtcaCAGAGGGgc	4	hsa-miR-2276-5p	
D	gTCACAGAggggc	2	hsa-miR-4677-3p	
D	gTCACAGAggggc	2	hsa-miR-4679	rs17227431
D	gtcacaGAGGGGC	4	hsa-miR-4749-3p	
C	GTCACAAaggggc	2	hsa-miR-6808-3p	
C	GTCACAAAggggc	2	hsa-miR-758-3p	
D	CTGGCCCtctgtg	4	hsa-miR-4640-5p	
D	CTGGCCCtctgtg	4	hsa-miR-4726-5p	rs17227424
D	cTGGCCCTctgtg	4	hsa-miR-6895-5p	
D	cCTGGCCCtctgt	4	hsa-miR-4640-5p	
D	cCTGGCCCtctgt	4	hsa-miR-4726-5p	
D	ccTGGCCCTctgt	4	hsa-miR-6895-5p	
C	cctGGCACTCtgt	4	hsa-miR-1909-5p	rs141300146
C	CCTGGCActctgt	8	hsa-miR-4755-3p	
C	cctggcACTCTGT	2	hsa-miR-4797-5p	
C	CCTGGCActctgt	8	hsa-miR-5006-5p	
D	ttcaCCGCCTGgc	2	hsa-miR-3937	
D	ttcaccGCCTGGC	10	hsa-miR-4755-3p	rs55830337

D	cttcaCCGCCTGg	2	hsa-miR-3937	rs56006529
C	cttcACTGCCTgg	2	hsa-miR-2682-5p	
C	cttCACTGCCtgg	2	hsa-miR-34a-5p	
C	cttcACTGCCTgg	2	hsa-miR-34b-5p	
C	cttCACTGCCtgg	2	hsa-miR-34c-5p	
C	cttCACTGCCtgg	2	hsa-miR-449a	
C	cttCACTGCCtgg	2	hsa-miR-449b-5p	
C	cttcACTGCCTgg	2	hsa-miR-449c-5p	
C	cttcaCTGCCTGg	2	hsa-miR-4514	
C	cttcacTGCCTGG	2	hsa-miR-4645-5p	
C	cttcacTGCCTGG	2	hsa-miR-4673	
C	cttcaCTGCCTGg	2	hsa-miR-4692	
				rs200745032
C	gggtTCAGGCTtc	4	hsa-miR-4293	

FuncClass :D: *The derived allele disrupts a conserved miRNA site (ancestral allele with support >= 2). C: The derived allele creates a new miRNA site*

*****